

# Wind Turbine Gearbox Vibration Signal Signature and Fault Development Through Time

Sofia Koukoura, James Carroll, Stepha Weiss and Alasdair McDonald

Department of Electronic and Electrical Engineering,

University of Strathclyde, Glasgow

Email: sofia.koukoura@strath.ac.uk

j.carroll@strath.ac.uk

stephan.weiss@strath.ac.uk

alsadair.mcdonald@strath.ac.uk

**Abstract**—This paper aims to present a methodology for health monitoring wind turbine gearboxes using vibration data. Monitoring of wind turbines is a crucial aspect of maintenance optimisation that is required for wind farms to remain sustainable and profitable. The proposed methodology performs spectral line analysis and extracts health features from harmonic vibration spectra, at various time instants prior to a gear tooth failure. For this, the tachometer signal of the shaft is used to reconstruct the signal in the angular domain. The diagnosis approach is applied to detect gear faults affecting the intermediate stage of the gearbox. The health features extracted show the gradient deterioration of the gear at progressive time instants before the catastrophic failure. A classification model is trained for fault recognition and prognosis of time before failure. The effectiveness of the proposed fault diagnostic and prognostic approach has been tested with industrial data. The above will lay the groundwork of a robust framework for the early automatic detection of emerging gearbox faults. This will lead to minimisation of wind turbine downtime and increased revenue through operational enhancement.

## I. INTRODUCTION

The challenges of climate change and energy security over the past years have increased the necessity to harvest wind energy. The cost of energy from wind farms is one of the main barriers that is currently restricting this renewable energy generation from being more widely adopted. Operation and maintenance costs constitute a large proportion of the total cost of energy from wind. Therefore, increased reliability plays a key role in reducing both those costs and the wind turbine downtime. Reliability can be increased through condition based maintenance, which is a form of preventative maintenance that involves continuous health monitoring of a wind turbine unit [1].

The wind turbine gearbox that bears alternating loads, is one of the most affected parts in a wind turbine, in terms of reliability [2]. Common failure modes include gear or bearing pitting, scuffing and fracture. If breakdown of the gearbox occurs it significantly increases the downtime and it is the most expensive component to maintain throughout the expected 20-year design life of a wind turbine. Effective health diagnosis of the gearbox is therefore vital in wind turbine fault detection and decision making.

Owing to the recent developments in the field of sensing and signal processing, modern wind turbines are equipped with condition monitoring systems for the on-line active remote monitoring of their components. Different type of sensory signals include vibration and electrical signals. Vibration analysis is one of the most commonly used mechanisms for condition monitoring of wind turbines. Based on the different vibration signatures, spectral line analysis methods measure the increase in the frequencies of the impulse signals and can indicate when a component failure is about to occur [3].

Vibration analysis is a particularly used to monitor the condition of gears. It requires the installation of acceleration transducers on the gearbox surface, which offers sensitivity in recognising faults. The main features that differentiate wind turbines from conventional rotary machines are the operating conditions under parameters of speed and variable load. Due to the rotor speed variation within the time window of the data acquisition, conventional signal processing techniques are insufficient because smearing into multiple bins can occur in the frequency domain. Therefore, computed order tracking is used and its principles are explained in [4]. An automated methodology that performs time synchronous averaging in a gearbox with limited speed variations that does not include a speed sensor is presented in [5]. Different vibration analysis methods are evaluated and presented in [6]. In terms of diagnosis methods, spectral kurtosis has been used to detect a tooth crack in the ring gear of a wind turbine planetary gearbox [7]. Sideband energy ratio can be calculated from spectrum data and it is sensitive to amplitude and frequency modulation, giving an indication of the health state of the gear [8]. Severity factors based on the sideband and kurtosis analysis are assigned in a diagnosis framework in [9]. As far as fault recognition via machine learning is concerned, an overview of Support Vector Machine (SVM) techniques for fault diagnosis and monitoring in engineering applications is presented in [10].

The above mentioned literature provides useful research approaches on wind turbine condition monitoring. However, the amount of data collected from the wind turbine and the complexity of this system call for an automated diagnostic process. Hence, this paper aims to present the groundwork

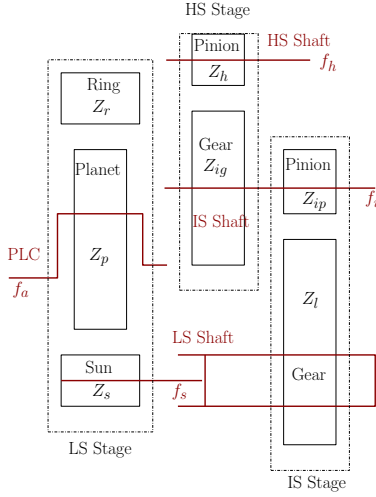


Fig. 1. Gearbox Internal Nomenclature

for the development an automatic gear fracture detection and prognostic methodology in a wind turbine gearbox, based on vibration signals. First of all, Section I introduces the papers motivation of research and refers to the research background which involves the analysis of the vibration signals and the fault detection process, taking into account the variable speed of the turbine. In Section II the structure of a wind turbine gearbox is explained and the proposed methodology is presented. Section III demonstrates the implementation of the methodology in real wind turbine vibration data. Section IV presents the results of the case studies and Section V concludes with the discussion and future work for the methodology development. The novelty in this paper is in the combination of identification techniques used to detect the failure and in the high quality data and failure example used to demonstrate the techniques in the case study.

## II. METHODOLOGY

### A. Wind Turbine Vibration Signal

1) *Wind Turbine Gearbox Structure*: A wind turbine gearbox is used to increase rotational speed from a low-speed rotor to a higher speed electrical generator. Since both compactness and high transmission ratio are desired, a typical wind turbine gearbox consists of three stages: one low speed planetary stage (PS) and two parallel stages, namely a high speed (HS) stage and an intermediate speed (IS) stage. The main shaft is connected to the planet carrier (PLC) and the HS pinion of the gearbox is geared to the generator.

The gearbox internal structure is shown in Figure 1 and the fundamental gear mesh frequencies (GMF) are determined in Table I.

2) *Wind Turbine Vibration Signals*: Vibration signals are obtained by acceleration transducers which are on the gearbox surface. Each transducer has different sensitivity and represents a different component of the gearbox. The multi-channel vibration signals are collected and analysed in the

TABLE I  
FUNDAMENTAL GEAR FREQUENCIES

Gear Element	# Teeth	Speed	GMF
PS Planet Carrier		$f_a$	
PS Ring Gear	$Z_r$	Fixed	
PS Planet Gear	$Z_p$	$f_a \frac{Z_r}{Z_p}$	$f_a Z_r$
PS Sun Pinion	$Z_s$	$f_s = f_a (1 - \frac{Z_r}{Z_s})$	
LS Gear	$Z_l$		
IS Pinion	$Z_{ip}$	$f_i = \frac{Z_l}{Z_{ip}} f_s$	$Z_{ip} f_i$
IS Gear Wheel	$Z_{ig}$		
HS Pinion	$Z_h$	$f_h = \frac{Z_{ig}}{Z_h} f_i$	$Z_h f_h$

frequency/order domain. The gear mesh frequencies are computed according to Table I.

### B. Angular Re-sampling

The rotational speed of the wind turbine rotor is determined by the torque controller in response to the wind speed and therefore is not constant. In variable-speed rotating machinery, vibration signals are non-stationary due to speed alteration. The frequencies of the vibrations though, are proportional to the rotational speed and the constants of proportionality are the orders. By re-sampling the non-stationary vibration signals, it is possible to reconstruct cyclo-stationary vibration signals in the angular domain, to avoid the mismatch between angle and time information.

The data is recorded by the tachometer and constant time increments. Each pulse of the tachometer signal represents a once-per-shaft-revolution event, so that is used to measure the shaft speed and is the reference for measuring the vibration phase angle. The signal is then up-sampled and low pass filtered. The re-sampled signal is interpolated linearly into a uniform phase domain grid. Then, the Short-Time Fourier Transform of the interpolated signal is computed.

### C. Fault Detection

When a gear has a local defect, such as a fatigue crack, the stiffness of the neighboring teeth is affected and this produces changes in the vibration signal. These changes are defined by amplitude and frequency modulation. The modulated gear meshing vibration is given by Eq. (1) [11].

$$y(t) = \sum_{m=0}^M X_m (1 + a_m(t)) \cos(2\pi m Z f_s t + \phi_m + \beta_m(t)) \quad (1)$$

Where  $Z$  is the number of teeth on the gear,  $f_s$  is the shaft rotational frequency and therefore  $Z f_s$  is the mesh frequency,  $a_m(t)$  is the amplitude modulation function,  $b_m(t)$  is the frequency modulation function,  $\phi_m$  is the initial phase of amplitude modulation and  $m$  is the integer.

As the modulation is periodic with the gear shaft rotation frequency  $f_s$ , these functions may be represented by discrete Fourier series, as in Eq. (2, 3).

$$a_m(t) = \sum_{n=0}^N A_{mn} \cos(2\pi n f_s t + a_{mn}) \quad (2)$$

$$b_m(t) = \sum_{n=0}^N B_{mn} \cos(2\pi n f_s t + \beta_{mn}) \quad (3)$$

Note that the modulation functions may differ with  $m$ .

In the frequency domain, the Fourier transform  $Y(f)$  will comprise the fundamental and harmonics of the meshing frequency surrounded by modulation sidebands. Multiple frequencies in the modulation cause multiple sidebands to appear in the spectrum. These sidebands occur at frequencies of  $Zf_s \pm kf_s$  where  $k$  is an integer of 1 or higher.

The signal analysis is performed through the MATLAB environment. Spectral line analysis is used to diagnose the health state of the vibration signals. After the application of order tracking, the spectrum has clear distinct order components which allows for an automatic procedure of algorithmic order peak detection. The diagnostic features used to analyse the health state of the vibration signals computed on the second harmonic of the gear mesh are the following:

- Sideband Energy Ratio (SER)
- Sideband Power Factor (SBPF)
- Sideband Average Power (SAP)
- Center Frequency and Sideband Kurtosis (KUR)
- Amplitude Ration (AR)

The SER algorithm sums the amplitudes of the first six sideband peaks on each side of the center mesh frequency  $\sum_{i=-6}^6 A_{SB,i}$  and divides by the amplitude of the center mesh frequency  $A_F$ , as in Eq. (4).

$$\text{SER} = \frac{\sum_{i=-6}^6 A_{SB,i}}{A_F} \quad (4)$$

For a healthy gear mesh the sidebands have a small amplitude compared to the center mesh frequency. As damage develops on a gear tooth, the sideband rise in amplitude which results in a larger SER value.

The SBPF algorithm sums the amplitudes of the center mesh frequency and the five sidebands on each side of the center mesh frequency, as in Eq. (5).

$$\text{SBPF} = A_F + \sum_{i=-5}^{+5} A_{(SB_i)} \quad (5)$$

The SAP algorithm calculates the total power of the first six sidebands rising around the center mesh frequency. The power is normalised with the length of the data segment.

$$\text{SAP} = \frac{\sum_{i=1}^N |A(i)|^2}{N} \quad (6)$$

Where  $N$  is the length of the narrowband data segment and  $A$  is the Amplitude of the signal.

Kurtosis is a measure of how outlier-prone a distribution is. Assuming that the part of the spectrum that includes the center mesh frequency and a frequency interval of up to six multiples of the shaft rotational speed represents a distribution, then the kurtosis is calculated. If the kurtosis has a high value it means that the distribution is quite sharp and most

values are concentrated in the center frequency. In case of a fault development, sidebands have increased amplitude in a spectrum, which means that the values are more distributed towards the tails, so the kurtosis is lower.

$$\text{KUR} = \frac{\frac{1}{N} \sum_{i=1}^N (A(i) - \bar{A})^4}{(\frac{1}{N} \sum_{i=1}^N (A(i) - \bar{A})^2)^2} \quad (7)$$

The amplitude ratio ensures that the center mesh frequency amplitude is higher than the maximum amplitude of the sideband, as shown in Eq.(8).

$$\text{AR} = \frac{A_F}{\text{Max}(A_{SB_1}, A_{SB_{-1}})} \quad (8)$$

If the sideband amplitude is higher than the center frequency, therefore  $\text{AR} < 1$ , it denotes a high severity level of fault.

#### D. Classification Training

Based on the features mentioned in the previous paragraph, a classification model is trained with the aim of fault recognition. It should be taken into account that the vibration signal from gears is affected greatly by the load. In order to distinguish such variations from changes in load, the reference torque is calculated based on the produced electrical power and generator. The extracted 5 vibration features along with the torque can be used to train the classification model.

However, as explained in [12], the number of available training samples should exceed the number of features, because the complexity of a model cannot exceed the complexity of the training dataset. Also often, the derived features are not linearly uncorrelated. In such cases, dimensionality reduction techniques, such as Principal Component Analysis (PCA) are applied [13].

Once a number of principal components are derived through PCA, these are used as inputs a classifier model that is expected to be able to distinguish between three classes of data: healthy, incipient fault and developed fault. For that purpose, a multi-class SVM is used. The exact mathematical formulation of SVM can be elaborated in [14]. The reason why SVM is preferred in this paper compared to other learning algorithms is due to its effectiveness when training small datasets [15].

### III. CASE STUDY

#### A. Gearbox Considered in this Study

The wind turbine considered in this study is rated at between 1.5 and 3 MW. The vibration data acquisition system consists of eleven accelerometers and a tachometer on the high speed shaft. The generator speed, the wind speed and the power produced by the turbine are also recorded. The acquisition time of the signal is between 10 and 11s and the sampling frequency is over 25kHz. Ranges are provided for the rated power, sampling period and sampling frequency for confidentiality reasons

The type of failure examined is a gear tooth failure. This failure occurs on the pinion tooth of the intermediate stage of the gearbox.

### B. Analysis of the Results

Data is collected for a wind turbine at various time steps prior to failure. The oldest dataset dates back to 2.5 years prior to failure and according to the maintenance reports the gearbox at this time is in a healthy state. The diagnosis framework is shown in Figure 2.

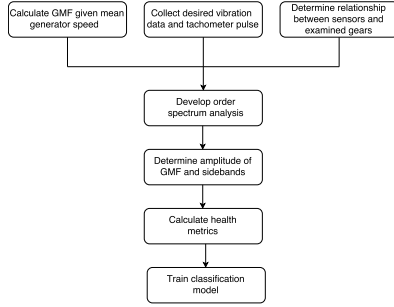


Fig. 2. Vibration Data Preprocessing and Diagnosis Procedure

## IV. RESULTS

### A. Case Study Results

The shaft speed variation for a given date prior to failure along with the tachometer pulse is shown in Figure 3. As depicted, the speed has a 5% variation during the 10.2s and therefore order tracking is applied as explained in Section II-B.

The vibration signatures from the sensor mounted on the intermediate speed shaft for similar loading conditions (Ref. Torque=0.8) is shown in Figure 4. The order is with respect to the high speed shaft, where the tachometer is mounted. The rising of sidebands around the second center mesh frequency (order 8.6) becomes more prominent in time steps closer to failure, as expected.

The health indicators as a function of torque are shown in Figure 5. The data is classified according to the time before failure that the signal was collected. Exponential fitted lines are also plotted for each time classification.

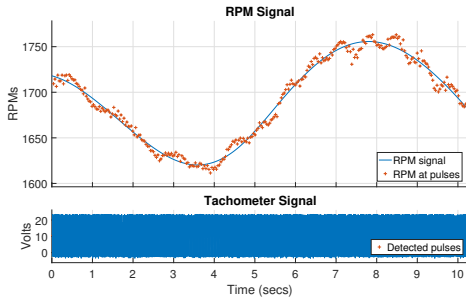


Fig. 3. Tachometer Pulse and RPM Signal

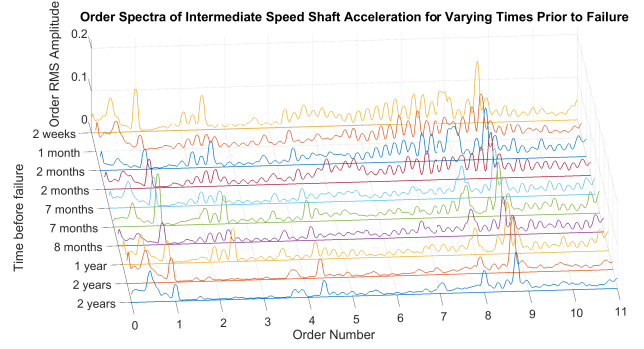


Fig. 4. Spectra Timeline Towards Failure

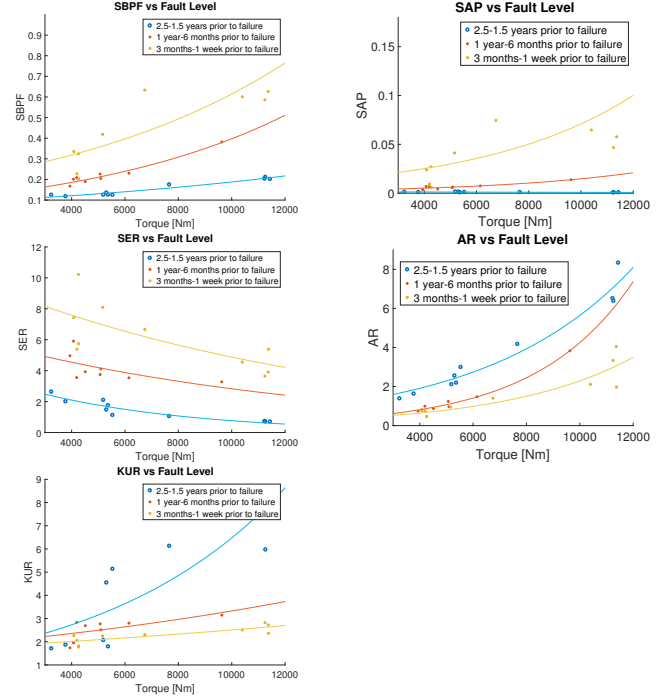


Fig. 5. Health Indicators as a Function of Torque

The SBPF and the SAP tends to increase with loading. Considering the healthy population, the values given from both algorithms are lower compared to the faulty population and in the case of the sideband energy they are almost zero. The SER algorithm also provides a promising indication of the fault, giving consistently higher values for faulty population compared to healthy, as expected based on the rising of sidebands.

The kurtosis provides an unclear diagnosis algorithm at low power levels, but as the power increases, the peakedness of the signal of the healthy population -and therefore the kurtosis- seems to increase considerably. On the other hand, the kurtosis of the faulty signal remains at lower levels as expected.

A 3-class classification model is trained based on the collection time of the signal relative to its failure. The classified data can be considered as 'healthy' for up to 1.5 years prior to failure, and according to the vibration spectra as

‘incipient fault’ for up to 6 months prior to failure and as ‘developed fault’ for the last 2 months prior to failure. A 5-fold cross validation method was used and the confusion matrix is shown in Figure 6. All fully developed fault data are classified correctly, which is the most crucial things in terms of maintenance, as a missed catastrophic failure can lead to large revenue loss and downtime. Regarding the incipient fault data, 2 out of 8 are misclassified; one as faulty and one as healthy. The first will lead to potentially earlier maintenance than scheduled and the second will lead to a missed detection which can however potentially be detected in later months. The misclassified healthy data can also lead to unscheduled maintenance.

Although the total accuracy of the classifier at 89.3% is fairly satisfactory, it should be noted that a larger dataset is needed in order to build a more robust model.

Confusion Matrix				
Predicted Class	developed fault	incipient fault	healthy	
	10 35.7%	0 0.0%	0 0.0%	100% 0.0%
	1 3.6%	6 21.4%	1 3.6%	75.0% 25.0%
	0 0.0%	1 3.6%	9 32.1%	90.0% 10.0%
				True Class
				developed fault   incipient fault   healthy
				90.9% 9.1%   85.7% 14.3%   90.0% 10.0%   89.3% 10.7%

Fig. 6. Confusion Matrix of Classification Model

### B. Validation of Vibration Analysis

This failure occurs on the pinion tooth of the intermediate stage of the gearbox. The tooth issue has been recorded as gear tooth tip and flank fracture, as shown in Figure 7. Root cause analysis of this gear tooth issue has not yet been completed by the OEM and is out with the scope of this paper.



Fig. 7. Broken Pinion on Intermediate Shaft

## V. CONCLUSIONS

This research showed that the developed framework for diagnosis of a gear fracture failure in a wind turbine gearbox

is quite effective, as demonstrated through a case study in real wind turbine data. For this, a healthy baseline of the gearbox is needed. The diagnosis of the gear state and the prognosis of time towards failure can be performed based on classification models. This methodology can potentially detect incipient faults almost 12 months prior to failure, which gives plenty of time to schedule maintenance.

In conclusion, future research steps include further development of the methodology, including broadband analysis health features, and development of threshold models. More case studies on the same type of failure mode need to be performed in order to develop a robust diagnostic model. Case studies in different types of gearboxes need to be implemented and the health features extracted will be used along with artificial intelligence techniques to predict failure of wind turbine components.

## REFERENCES

- [1] C. J. Crabtree, D. Zappalá, and S. I. Hogg, “Wind energy: Uk experiences and offshore operational challenges,” *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, vol. 229, no. 7, pp. 727–746, 2015.
- [2] J. Carroll, A. McDonald, and D. McMillan, “Failure rate, repair time and unscheduled o&m cost analysis of offshore wind turbines,” *Wind Energy*, 2015.
- [3] W. Lu and F. Chu, “Condition monitoring and fault diagnostics of wind turbines,” in *Prognostics and Health Management Conference, 2010. PHM’10*. IEEE, 2010, pp. 1–11.
- [4] K. Fyfe and E. Munck, “Analysis of computed order tracking,” *Mechanical Systems and Signal Processing*, vol. 11, no. 2, pp. 187–205, 1997.
- [5] F. Combet and L. Gelman, “An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor,” *Mechanical systems and signal processing*, vol. 21, no. 6, pp. 2590–2606, 2007.
- [6] S. Sheng, “Wind turbine gearbox condition monitoring round robin study–vibration analysis,” *Contract*, vol. 303, pp. 275–3000, 2012.
- [7] T. Barszcz and R. B. Randall, “Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine,” *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 1352–1365, 2009.
- [8] J. Hanna, C. Hatch, M. Kalb, A. Weiss, and H. Luo, “Detection of wind turbine gear tooth defects using sideband energy ratio,” *China Wind Power 2011; October, 19-21, 2011, Beijing, China*, 2011.
- [9] P. Tamilselvan, P. Wang, S. Sheng, and J. M. Twomey, “A two-stage diagnosis framework for wind turbine gearbox condition monitoring,” *International Journal of Prognostics and Health Management*, pp. 21–31, 2013.
- [10] V. Khachatryan, A. Sirunyan, A. Tumasyan, W. Adam, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö *et al.*, “Search for resonant  $t\bar{t}$  production in proton-proton collisions at  $\sqrt{s} = 8$  tev,” *Physical Review D*, vol. 93, no. 1, p. 012001, 2016.
- [11] P. McFadden, “Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration,” *Journal of vibration, acoustics, stress, and reliability in design*, vol. 108, no. 2, pp. 165–170, 1986.
- [12] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty, “Optimal number of features as a function of sample size for various classification rules,” *Bioinformatics*, vol. 21, no. 8, pp. 1509–1515, 2004.
- [13] C. Skitides and W.-G. Früh, “Wind forecasting using principal component analysis,” *Renewable Energy*, vol. 69, pp. 365–374, 2014.
- [14] V. N. Vapnik, “An overview of statistical learning theory,” *IEEE transactions on neural networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [15] Z. Yin and J. Hou, “Recent advances on svm based fault diagnosis and process monitoring in complicated industrial processes,” *Neurocomputing*, vol. 174, pp. 643–650, 2016.